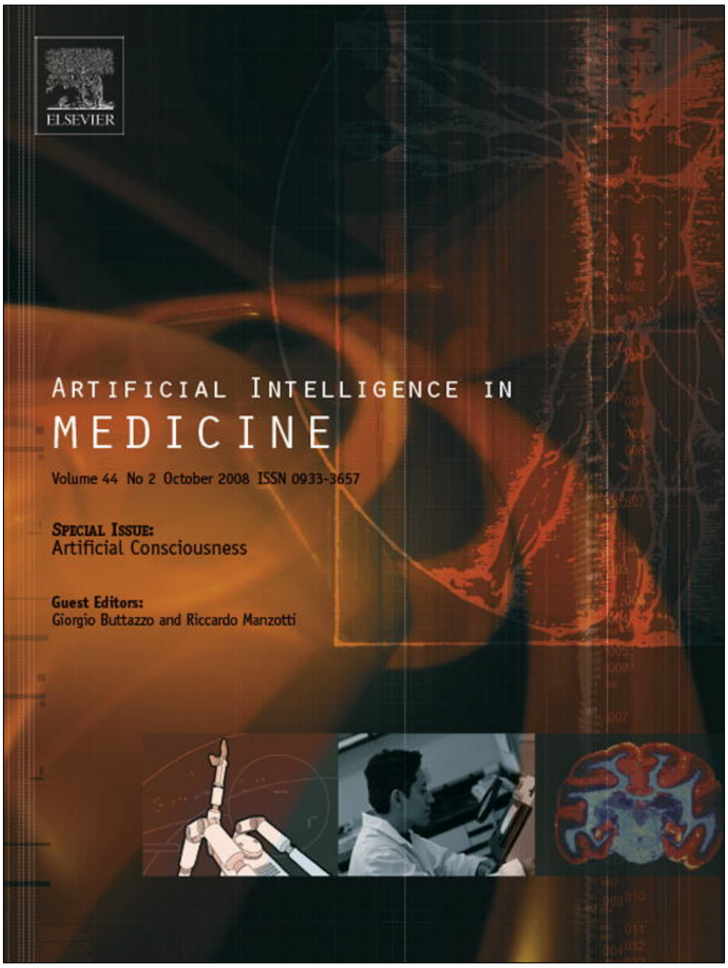
ELSEVIER
ARTIFICIAL INTELLIGENCE IN
MEDICINE
Volume 44 No 2 October 2008 ISSN 0933-3657
SPECIAL ISSUE:
Artificial Consciousness
Guest Editors:
Giorgio Buttazzo and Riccardo Manzotti

Artificial Intelligence in Medicine (2008) 44, 139—146

ARTIFICIAL INTELLIGENCE IN MEDICINE

http://www.intl.elsevierhealth.com/journals/aiim

Artificial consciousness: Hazardous questions (and answers)

Giorgio Buttazzo *

Scuola Superiore Sant'Anna, Via Moruzzi 1, 56124 Pisa, Italy

Received 19 February 2008; received in revised form 30 June 2008; accepted 3 July 2008

KEYWORDS
Artificial consciousness;
Conscious machines;
Artificial neural networks;
Artificial beings;
Self-aware systems

Summary If human consciousness is the result of complex neural electro-chemical interactions occurring in the brain, the question of whether a machine can ever become self-aware could be a matter of time: the time necessary to fully understand the functional behavior of the brain structure, develop a mathematical model of it, and implement an artificial system capable of working according to such a model.

This paper addresses several issues related to the possibility of developing a conscious artificial brain. A number of hazardous questions are posed to the reader, each addressing a specific technical or philosophical issue, which is discussed and developed in a form of a hazardous answer.

1. Introduction

The idea of building a smart machine capable of competing with human intelligence has always been a dream of some computer scientists since the early establishment of computer technology [1,2]. Today, given the successful results of artificial intelligence and neural computing, and considering the current pace of computer evolution, such a possibility seems to be more concrete than ever, and someone believes that, in the near future, machines will exceed human intelligence and eventually will develop a mind. Talking about artificial consciousness, however, gives rise to many philosophical issues [3]. Are computers thinking, or are they just calculating? Conversely, are human beings thinking, or are they just calculating? Is consciousness a prerogative of human beings? Does it depend on the material the brain is made of or can it be replicated using a different hardware? Answering these questions is not easy, since it requires moving along the edges of several different disciplines, such as computer science, neurophysiology, philosophy, and religion. Nevertheless, many people believe that artificial consciousness is possible and that in the future it will emerge in complex computing machines. In the rest of this paper, a number of provocative questions are posed to the reader, each addressing a specific technical or philosophical issue, which is discussed and developed in a form of a hazardous answer.

2. What is artificial consciousness?

As Tagliasco pointed out [4], the term "artificial" is often used in two different meanings. In a first form,

* Tel.: +39 050 882012; fax: +39 050 882003.
E-mail address: buttazzo@sssup.it.

0933-3657/$ — see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.artmed.2008.07.004

the term "artificial" denotes a human artifact that replicates or simulates a real thing. For example, an artificial flower is something that appears as a flower, may have the same shape and colors, but it is very different in terms of materials and functions. In this sense, an artificial thing is a very simplistic version of its real counterpart. In some other cases, the term "artificial" is used to denote a genuine physical phenomenon reproduced using a human made device. For example, "artificial light" denotes an electromagnetic wave produced by a human made device, like a bulb or a led. Depending on the meaning we associate with the term "artificial", we can distinguish two types of artificial consciousness, as proposed by Holland [5,6]:

- *Weak artificial consciousness*: It is a simulation of a conscious behavior. It can be implemented as a smart program that simulates the behavior of a conscious being at some level of detail, without understanding the mechanisms that generate consciousness.
- *Strong artificial consciousness*: It refers to a real conscious thinking emerging from a complex computing machine (artificial brain). In this case, the main difference with respect to the natural counterpart depends on the hardware that generates the process.

In this paper we are interested in Strong Artificial Consciousness and we will speculate on the possibility that such a form of consciousness can emerge in a complex computing system.

3. How can we verify consciousness?

In 1950, the computer science pioneer Alan Turing posed a similar problem but concerning intelligence. In order to establish whether a machine can or cannot be considered intelligent as a human, he proposed a famous test, known as the Turing test: there are two keyboards, one connected to a computer, the other leads to a person. An examiner types in questions on any topic he likes; both the computer and the human type back responses that the examiner reads on the respective computer screen. If the examiner cannot reliably determine which was the person and which the machine, then we say the machine has passed the Turing test.

In 1990, the Turing Test received its first formal acknowledgement from Hugh Loebner (a New York philanthropist) and the Cambridge Center for Behavioral Studies (Cambridge, MA), which established the Loebner Prize Competition in Artificial Intelligence [7]. Loebner pledged a prize of $100,000 for the first computer whose responses were indistinguishable from those of a human. The first competition was held at the Computer Museum of Boston in November 1991. For some years, the contest was constrained to a single narrow topic, but the most recent competitions, since 1998, did not limit the scope of questioning. Each judge, after the conversation, gives a score from 1 to 10 to evaluate the interlocutor, where 1 means human and 10 computer. So far, no computer has given responses totally indistinguishable from a human, but every year scores are getting closer to five in the average [8]. Today, the Turing test can be passed by a computer only if we restrict the interaction on very specific topics, as chess.

On 11 May 1997 (3:00 p.m. eastern time), for the first time in the history, a computer named Deep Blue beat world chess champion Garry Kasparov, 3.5—2.5. As all actual computers, however, Deep Blue does not understand chess, since it just applies some rules to find a move that leads to a better position, according to an evaluation criterion programmed by chess experts.

Claude Shannon estimated that in a chess game the search space includes about 10^{120} possible positions. Deep Blue was able to analyze 200 million (2×10^8) positions per second. Exploring the entire search space for Deep Blue would therefore take about 5×10^{111} s, which is about 10^{95} billions of years. Nevertheless, Deep Blue victory can be attributable to its speed combined with a smart search algorithm, able to account for positional and material advantage. In other words, computer superiority was due to brute force, rather than sophisticated machine intelligence.

In spite of that, in many interviews during and after the match, Kasparov expressed doubts he was playing against the computer and sometime he felt like playing against a human. In some situation, he also appreciated the beauty of the moves done by the machine, as if it was driven by intention, rather than by brute force. Thus, if we accept Turing's view, we can say that Deep Blue plays chess in an intelligent way, but we can also claim that it does not understand the meaning of his moves, as a television set does not understand the meaning of the images it displays.

Besides chess, there are other domains in which computers are reaching human ability, and their number is increasing every year. In music, for example, there are many commercial programs that can create melodic lines or even entire songs according to specific styles, ranging from Bach to jazz. There are also programs that generate great solos on top of a given chord sequence, emulating jazz masters, like Charlie Parker and Miles Davis, much better than an average human musician can do. In 1997, Steve Larson, a music professor at University of Oregon, proposed a musical variation of the Turing test by asking an audience to listen to a set of pieces of classical music and determine which one was written by a computer and which was the authentic composition. As a result, many computer pieces were classified as authentic compositions and vice versa, meaning the computer passed the Turing test on music (see [9], page 160).

Other areas where computers are becoming as smart as humans include continuous speech understanding, electrocardiogram diagnostics, theorem proving, and aircraft guidance. In the next future, such domains will quickly expand to include more complex tasks such as car driving, real-time language translation, house cleaning, medical surgery, surveillance, law enforcement, and so on.

But if machines will become as good as humans in many disciplines, such that they are indistinguishable from them in the sense of the Turing test, does it mean they are self-aware? Of course, no. However, determining self-awareness is a more delicate issue. In fact, if intelligence is an expression of an external behavior that can be measured by specific tests, self-consciousness is a property of an internal brain state, which cannot be measured. Hence, to address the issue we are forced to make some philosophical consideration.

From a pure philosophical point of view, it is not possible to verify the presence of consciousness in another brain (either human or artificial), because this is a property that can only be verified by his possessor. Since we cannot enter in another being's mind, then we cannot be sure about his consciousness. Such a problem is deeply discussed by Hofstadter and Dennett, in a book entitled The Mind's I [10].

From a pragmatic point of view, however, we could follow Turing's approach and say that a being can be considered self-conscious if able to convince us, by passing specific tests. Moreover, among humans, the belief that another person is self-conscious is also based on similarity considerations: since we have the same organs and we have a similar brain, it is reasonable to believe that the person in front of us is also self-conscious. Who would question his best friend's consciousness? Nevertheless, if the creature in front of us, although behaving like a human, were made by synthetic tissues, mechatronic organs, and neural processors, our conclusion would be perhaps different.

The most common objection to this issue is that computers, being driven by electronic circuits working in a fully automated mode, cannot exhibit creativity, emotions, love, or free will. A computer is a slave operated by its components, just as a washing machine.

4. Can ever computers think?

In 1980, the philosopher Searle [11] claimed to be able to prove that no computer program could possibly think or understand, independently of its complexity. His proof is based on the fact that every operation that a computer is able to carry out can equally well be performed by a human being working with paper and pencil in a disciplined but unintelligent manner.

Searle's Chinese room objection considers a situation in which, at some stage in the future, a computer is able to pass a Turing test conducted in Chinese, so being indistinguishable from human beings. Searle claims that, no matter how good the performance of the program, it cannot in fact think and understand. This can be proved, he says, by considering an imaginary human being, who speaks no Chinese, hand working the program in a closed room. The interogator's questions, expressed in the form of Chinese ideograms, enter the room through an input slot. The human in the room follows the instructions in the program and carries out exactly the same series of computations that an electronic computer running the program would carry out. These computations eventually produce strings of ideograms through an output slot. As far as the waiting interogator is concerned, the ideograms form an intelligent response to the question that was posed. But, as far as the human in the room is concerned, the output is completely meaningless. He does not even know that the inputs and outputs are linguistic expressions, nor does an electronic computer.

The problem with this reasoning is that it also applies to the biological counterpart. In fact, at a neural level, human brain is also operated by electrochemical reactions and each neuron automatically responds to its inputs according to fixed laws. Each neuron is not conscious, but contributes to thinking without understanding what is going on in our mind. However, this does not prevent us of experiencing happiness, love and irrational behaviors.

With the emergence of artificial neural networks, the problem of artificial consciousness becomes even more intriguing, because neural networks replicate the basic electrical behavior of the brain and provide the proper support for realizing a processing mechanism similar to the one adopted by the brain. In the book "Impossible Minds", Aleksander [12] addresses this topic with depth and scientific rigor.

If we remove structural diversity between biological and artificial brains, the issue about artificial consciousness can only become religious. In other words, if we believe that human consciousness is determined by divine intervention, then clearly no artificial system can ever become self-aware. If instead we believe that human consciousness is a natural property developed by complex brains, then the possibility of realizing an artificial self-aware being remains open.

5. Is consciousness separated from the brain?

From a religious point of view, the main argument against the possibility of replicating self-awareness in an artificial system is that consciousness is not a product of the brain activity, but a separate immaterial entity, often identified with the soul. Such a dualistic theory about brain and mind was mainly developed by Rene Descartes (1596—1650) and is still shared by many people. However, it lost credibility in the philosophical community, since it encounters several problems that cannot be explained with it.

1. First of all, if a mind is separated from its brain, how can it physically interact with the body and activate a neural circuit? Whenever I think to move, specific electrochemical reactions take place in my brain, which make some neurons start firing to actuate the desired muscles. But, if mind operates outside the brain, how is it able to move atoms to create electrical pulses? Are there mysterious forces that activate neural cells? Does a mind interact with a brain by violating the fundamental laws of physics?
2. Second, if a conscious mind can exist outside the brain, why do we have a brain?
3. Third, if emotions and thoughts come from outside, why does a brain stimulated with electrodes and drugs responds by generating thoughts?
4. Finally, why in patients with brain diseases conscious behavior is severely affected by surgically removing portions of their brain?

These and other arguments caused dualism to loose credibility in the scientific and philosophical community. To solve such inconsistencies, many other alternative formulations were developed to address the mind/brain issue. From one side, reductionism did not recognize the existence of mind as a private sense data and considered all mental activities as specific neural states of the brain. On the other hand, idealism tried to refuse the physical world by considering all events as mental constructions. Unfortunately, here there is not space to discuss the various theories on the subject, and this is out of the scope of this article. However, it is important to point out that, with the progress of Computer Science and Artificial Intelligence, a new approach took root among scientists and philosophers, according to which mind is considered as a form of computation emerging at a higher level of abstraction with respect to neural activity.

The major weakness of the reductionist approach in the comprehension of mind is to recursively decompose a complex system into simpler subsystems, until at some stage the units can be fully analyzed and described. This method works perfectly for linear systems, where any output can be seen as a sum of simpler components. However, a complex system is often non-linear, thus analyzing its basic components it is not sufficient to understand its global behavior. In such systems, there are *holistic* features that cannot be seen at a smaller level of detail, but they appear only when considering the structure and the interactions among the components.

Davies, in his book *God and the New Physics* [13], explains this concept by observing that a digital picture of a face consists of a large number of colored dots (pixels), each of which does not represent the face: the shape takes his form only when we observe the picture at a certain distance which allows us to see all the pixels. The face is not a property of the pixels as such, but of the set of pixels.

In *Gödel, Escher, Bach* [14], Hofstadter explains the same concept by describing the behavior of a large ant colony. As known, ants have a complex and highly organized social structure based on work distribution and collective responsibility. Although each ant has a very little intelligence and limited capabilities, the whole ant colony exhibits a highly complex behavior. In fact, building an ant nest requires a large and complex design, but clearly, no individual ant has in mind a complete picture of the whole project. Nevertheless, a scheme and a finalized behavior emerge at the colony level. In some sense, the whole colony can be considered as a living being.

From many aspects, a brain is similar to a large ant colony, since it consists of billions of neurons that cooperate for achieving a common objective. Interaction among neurons is much tighter than among ants, but the underlying principles are similar: work subdivision and collective responsibility. Consciousness in not a property of individual neurons, which just automatically operates as switches, responding to input signals. Consciousness is rather a holistic property that emerges and flourishes from neural cooperation when the system reaches a sufficiently organized complexity.

Although most people have no problem of accepting the existence of holistic features, someone still believes that consciousness cannot emerge from a silicon substratum, being an intrinsic property of biological materials, like neural cells. So it is reasonable to ask the following question.

6. Does consciousness depend on the material neurons are made of?

Paul and Cox, in *Beyond Humanity* [15] say:

"It would be astonishing that the most powerful information-processing tool can be derived only from organic cells and chemistry. Aircraft is made out of different materials from birds, bats, and bugs; solar panels are made out of different materials from leaves. There is usually more than one way to build a given type of machine. [...] This just happens to be what genetics was able to work with. [...] Other elements and combinations of elements may be able to do a better job processing information in a self-aware manner."

If we support the hypothesis of consciousness as a physical property of the brain, then it is reasonable to ask "when" a machine will become self-aware.

7. When will a machine become self-aware?

Attempting to provide even a rough answer to this question is hazardous. Nevertheless, it is possible to determine at least a necessary condition, without which a machine cannot develop self-awareness. The idea is based on the simple consideration that, to develop self-awareness, a neural network must be at least as complex as the human brain.

This is a reasonable assumption to start with, since it seems that less complex brains are not able to produce conscious thoughts. Consciousness seems to be a step function of brain complexity, where the threshold is the one of human brain (see Fig. 1).

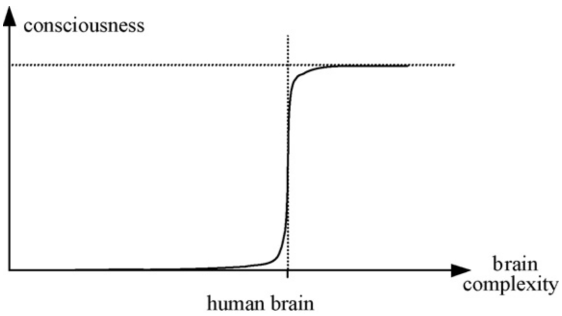

Figure 1 Consciousness as a function of brain complexity.

But how complex is the human brain? How much memory is required to simulate its behavior with a computer? The human brain has about 10^{12} neurons, and each neuron makes about 10^3 connections (synapses) with other neurons, in the average, for a total number of 10^{15} synapses. In artificial neural networks, a synapse can be simulated using a floating-point number requiring 4 bytes of memory to be represented in a computer. As a consequence, to simulate 10^{15} synapses a total amount of 4×10^{15} bytes (4 millions of Gigabytes) is required. Let us say that to simulate the whole human brain we need 5 millions of Gigabytes, including the auxiliary variables for storing neuron outputs and other internal brain states. Then, when will such a memory be available in a computer?

During the last 20 years, the RAM capacity increased exponentially by a factor of 10 every 4 years. The plot in Fig. 2 illustrates the typical memory configuration installed on personal computers since 1980.

By interpolation, we can derive the following equation, which gives the RAM size (in bytes) as a function of the year:

$$\text{bytes} = 10^{((\text{year}-1966)/4)}$$

For example, from the equation we can find that in 1990 a personal computer was typically equipped with 1 Mbytes of RAM. In 1998, a typical configuration had 100 Mbytes of RAM, and so on. By inverting the relation above, we can predict the year in which a computer will be equipped with a given amount of memory (assuming the RAM will continue to grow at the same rate):

$$\text{year} = 1966 + 4\log_{10}(\text{bytes}).$$

Now, to know the year in which a computer will be equipped with 5 millions of Gbytes of RAM, we have just to substitute that number in the equation above and compute the result. The answer is:

$$\text{year} = 2029.$$

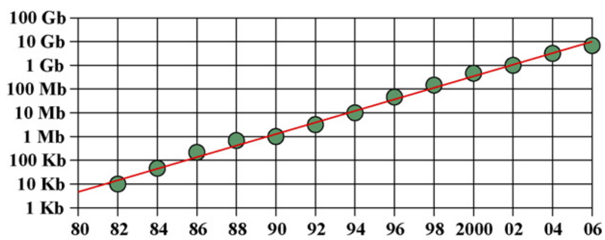

Figure 2 Typical RAM configurations (in bytes) installed in personal computers in the last 20 years.

An interesting coincidence with the date predicted in Terminator's movie [16]. It is also worth observing that similar predictions were derived by Moravec [17], Kurzweil [9], Paul and Cox [15].

In order to fully understand the meaning of the achieved result, it is important to make some considerations. First of all, it is worth recalling that the computed date only refers to a necessary, but not sufficient, condition to the development of an artificial consciousness. This means that the existence of a powerful computer equipped with millions of Gigabytes of RAM is not sufficient alone to guarantee that it will magically become self-aware. There are other important factors influencing this process, such as the progress of theories on artificial neural networks and on the basic biological mechanisms of mind, for which is impossible to attempt precise estimates. Furthermore, someone could argue that the presented computation was done on personal computers, which do not represent the top of technology in the field. Some other could object that the same amount of RAM memory could be available using a network of computers or virtual memory management mechanisms to exploit hard disk space. In any case, even if we adopt different numbers, the basic principle of the computation is the same, and the date could be advanced by a few years only.

Someone may object that the prediction of the 2029 date relies on mindless extrapolation of current trends, without considering events that may alter that trend. Indeed, the exponential growth of computing power and memory was noted in 1973 by Gordon Moore, one of Intel's founders, who predicted that *the number of transistors on integrated circuits would continue to double every 18 months until fundamental physical limits are reached*. The accuracy of the prediction over the past 25 years was such that it is referred to as "Moore's Law". But how much longer this law will continue to hold true in the future? Chip companies estimated that Moore's Law will continue to be valid for another 15 or 20 years. Then, when transistors will reach the size of a few atoms, the conventional approach will not work and this paradigm will break down. What next? Will microprocessor evolution come to the end around the year 2020?

Some people, including Kurzweil [9] and Moravec [17], noticed that computers have been growing exponentially in power long before the invention of the integrated circuit in 1958, regardless of the type of hardware used. So Moore's Law on integrated circuits was not the first, but the fifth paradigm to continue the exponential growth of computing. Each new paradigm came along just when needed. This suggests that exponential growth will not stop with the end of Moore's Law. The industry is not without choices for the future and new technologies are being investigated by scientists, such as three-dimensional chip design, optical computing, and quantum computing [18]. Thus, although Moore's Law will not hold in the future (since it does not apply to non-silicon-based systems), the exponential growth of computing power will probably continue for many years to come.

8. Can a sequential software program become self-conscious?

If consciousness is a product of a highly organized information processing system, then it does not depend on the hardware substratum, but on the information processed by that hardware. Then simulating the hardware with a software program would produce the same result. Indeed, most artificial neural networks today are simulated by sequential programs running on a single processor. A simulation software is much more flexible (although slower) than a hardwired network, since it allows changing simulation parameters, as well as the network architecture, in a more flexible way. One could argue that a simulation of a process is different than the process itself. Clearly, this is true when simulating a physical phenomenon, like a thunderstorm or a planetary system. However, for a neural network, a simulation is not different from the process, because both are information processing systems. Similarly, the software calculator available in most PC operating systems performs the same operations as its hardware counterpart: they are functionally equivalent. Hence, if we accept that consciousness is the product of a complex information processing system, we must accept that artificial consciousness (in the strong sense) can also be generated by sequential simulation software.

9. How would we feel with a faster brain?

Does consciousness depend on the speed of the computing elements? It is hard to say, but intuition suggests that it should be independent, since the results of a computation do not depend on the hardware where the process is performed. However, processing speed is important to meet the real-time requirements imposed by the external world. If we could ideally slow down our neurons uniformly in the whole brain, we would perhaps perceive the world as a fast-motion movie, where events occur faster than our reactive capabilities. That sounds reasonable, because our brain evolved and adapted in a world where the important events for reproduction and survival are within a few tenths of a second scale. If we could speed up the events in the environment or we could slow down our neurons, we would not be able to operate in real-time in such a world any more, and probably we would not survive.

Conversely, how would we feel having a faster brain? Today a logic port is six orders of magnitude faster than a neuron. While biological neurons respond within a few milliseconds (10^{-3} s), electronic circuits respond within a few nanoseconds (10^{-9} s). This observation leads to an interesting question: if consciousness will emerge in an artificial machine, what will time perception be like to a brain that thinks millions of times faster than a human brain?

It is possible that for conscious machines the world around them seems to move slower. Perhaps the same thing happens to insects, having a smaller but fast reactive brain. Perhaps, to a fly, a human hand that tries to swat looks like it is moving in slow motion, giving the fly plenty of time to glide leisurely out of the way.

Paul and Cox address this issue in *Beyond Humanity* [15]:

"... a cyberbeing will be able to learn and think at hyperspeed. They will observe a speeding bullet fired from a moderate distance away, calculate its trajectory, and dodge it if necessary. [...] Imagine being a robot near a window. To your fast thinking mind, a bird takes what seems like hours to cross the field of view, and a day lasts seemingly forever."

It is worth noticing that the issue of time perception has been somehow addressed in *The Matrix* movie (Larry and Andy Wachowski, 1999).

10. Are there different levels of consciousness?

There is strong evidence that consciousness is a process that emerges in a highly interconnected network of neural cells, each of which is not a conscious entity. Like an ant, which blindly operates according to its instinctive instructions coded in its DNA, each neuron just follows the laws of physics and it is unaware of what is going on in the whole brain. Similarly, we could ask:

"Could some form of consciousness emerge in a large geographical network of unconscious computers, as a consequence of their complex interactions?"

Or:

"Could human beings be responsible for generating a new form of conscious thinking as a consequence of their social interactions at a planetary level?"

Looking at lights distribution in a landscape from a plane during a night flight, it is interesting to observe how cities tend to spatially expand following a highly branched structure similar to neural cells with lot of dendrites, where roads represent neural connections and moving vehicles represent elements for exchanging information. Is human society unconsciously weaving a sort of neural system around the Earth from which a new form of conscious thinking could emerge as a consequence of the complex interactions of its elements?

If such a conjecture is correct, two observations are still worth making. First of all, if this is going to happen sometime in the future, it will be extremely difficult, if not impossible, to verify. In fact, such a new form of intelligence would operate with a different time scale, several orders of magnitude slower than our thinking speed. Considering the time taken by humans to travel between sites, a simple conscious thought of such a hyper-being would take days or even months. Similarly, conscious thinking in our brain (evolving in the order of seconds) evolves thousand times slower than neuron reaction times (milliseconds). Space and time are often related in nature: the bigger a system, the slower it evolves. The second observation is that this sort of neural system is self-repairing. In fact, if a catastrophic event destroys a communication channel (e.g., a road), special active elements (human workers) would cooperate to repair the damage, exactly like blood cells contribute to repair blood vessels. Like a biological neural system, the hyper-being would also be self-adaptive, since intensive communication between two units (e.g., cities) would reinforces the channel, stimulating the development of new roads between them to satisfy the need of an increased traffic.

11. Conclusions

In this paper we addressed several issues related to the possibility of developing a conscious artificial brain. Assuming that human consciousness is the result of complex neural electro-chemical interactions, the possibility that machines will exceed human intelligence and eventually will develop a mind is becoming more realistic. Considering the current pace of computers evolution and the progress of artificial neural networks, many scientists predicted that computing systems will reach the complexity of the human brain around 2030. Such a point in time is what futurists call "the technological singularity" [19], an event that could revolutionize our life. Indeed, the creation of artificial conscious machines that would quickly become smarter than humans would have enormous implications in human race evolution [20].

But "*why developing a self-aware machine*?" Except for ethical issues, that could significantly influence the progress in this field, the strongest motivation for developing a conscious machine would certainly come from the innate human desire of discovering new horizons and enlarging the frontiers of science. And this makes this process to be unavoidable. One of the major implications is that, developing an artificial brain based on the same principles used in the biological brain, would provide a way for transferring our mind into a faster and more robust support, opening a door towards immortality. Freed from a fragile and degradable body, human beings with synthetic organs (including brain) could represent the next evolutionary step of human race. Such a new species, natural result of human technological progress, could start the exploration of the universe, search for alien civilizations, survive to the death of the solar system, control the energy of black holes, and move at the speed of light by transmitting the information necessary for replication on other planets.

Indeed, the exploration of space aimed at searching for extraterrestrial intelligent civilizations already started in 1972, when Pioneer 10 spacecraft was launched to go out of our solar system with the specific purpose of broadcasting information about human race and planet Earth in the open space, as a bottle in the ocean. As for all important human discoveries, from nuclear energy to atomic bomb, from genetic engineering to human cloning, the real problem has been and will be to keep technology under control, making sure that it is used for human progress, not for catastrophic aims.

References

[1] Asimov I. "I, Robot", a collection of short stories originally published between 1940 and 1950. London, UK: Grafton Books; 1968.

[2] Stork DG (ed). "HAL's legacy: 2001's computer as dream and reality" edited by David G. Stork, Foreword by Arthur C. Clarke, MIT Press, Cambridge, MA, 1997.

[3] Buttazzo G. Can a machine ever become self-aware? In: Aurich R, Jacobsen W, Jatho G, editors. Artificial Humans: An Historical Retrospective of the Berlin International Film Festival 2000. Los Angeles: Goethe Institute; May 2000. p. 45—9.

[4] Tagliasco V. Artificial consciousness: a technology discipline. In: Chella A, Manzotti R, editors. Artificial Consciousness. Exeter, UK: Imprint Academic; January 2007.

[5] Holland O. Machine consciousness. New York, NY: Imprint Academic; 2003.

[6] Manzotti R. From artificial intelligence to artificial consciousness. In: Chella A, Manzotti R, editors. Artificial consciousness. London, UK: Imprint Academic; January 2007

[7] Home page of the Loebner Prize, 1999: http://www.loebner.net/Prizef/loebner-prize.html (Accessed 25 June 2008).

[8] Krol M. Have we witnesses a real-life Turing test? IEEE Comput March 1999;32(3):27—30.

[9] Kurzweil R. The age of spiritual machines. New York, NY: Viking Press; 1999.

[10] Hofstadter DR, Dennett DC. The mind's I. New York, NY: Harvester/Basic Books; 1981.

[11] Searle J. Minds, brains, and programs. Behav Brain Sci 1980;3:417—24.

[12] Aleksander I. Impossible minds, my neurons my consciousness. London, UK: Imperial College Press; 1996.

[13] Davies PC. God and the new physics. New York, NY: Simon and Schuster Trade; 1983.

[14] Hofstadter D, Gödel, Escher, Bach. Basic books, New York, NY, 1979.

[15] Paul GS, Cox E. Beyond humanity: cyberevolution and future minds. Rockland, MA, USA: Charles River Media, Inc.; 1996.

[16] Cameron J, Wisher W. "Terminator 2: judgment day" (1991), the movie script, http://www.scifiscripts.com/scripts/t2.txt (Accessed: 25 June 2008).

[17] Moravec H. Robot: mere machine to transcendent mind. New York, NY: Oxford University Press; 1999.

[18] Geppert L. Quantum transistors: toward nanoelectronics. IEEE Spect September 2000;37(9):46—51.

[19] Horgan J. The consciousness conundrum. IEEE Spect June 2008;45(6):28—33.

[20] Vinge V. Signs of the singularity. IEEE Spect June 2008;45(6): 69—74.